

CLAIMS

1. A data-consistent, distributed unified data set device group that stores a distributed unified data set, the distributed unified data set device group comprising:

5 a first data storage component, within a first supervisor local data-storage device, that stores a first portion of the distributed unified data set, the first supervisor local data-storage device receiving WRITE requests directed to the first portion of the distributed unified data set from a host computer via a communications medium;

10 a second data storage component, within a second subordinate local data-storage device, that stores a second portion of the distributed unified data, the second subordinate local data-storage device ; and

15 a mirror unified data set that mirrors the distributed unified data set stored on one or more remote data-storage devices.

2. The data-consistent, distributed unified data set device group of claim 1 wherein the first supervisor local data-storage device includes a controller that executes a unified sequence number component that provides sequence numbers for WRITE requests directed to the data-consistent, distributed unified data set device group and that executes controller functionality for receiving, carrying out, and responding to WRITE requests received from a host computer, and wherein the second subordinate local data-storage device includes a controller that executes controller functionality for receiving, carrying out, and responding to WRITE requests received from a host computer.

3. The data-consistent, distributed unified data set of claim 2 wherein the controller functionality of a local data-storage device for receiving, carrying out, and responding to WRITE requests received from a host computer, upon receiving a WRITE request directed to a portion of the distributed unified data set contained within a data-storage component of the local data-storage device, transmits a mirror WRITE request to the one or more remote data-storage devices, including in the mirror WRITE a unified sequence requested by the controller functionality from the unified-sequence-number component executed by the controller of the supervisor local data-storage device.

4. The data-consistent, distributed unified data set of claim 3 wherein, when the
unified-sequence-number component of the supervisor local data-storage device receives a
request for a unified sequence number to include in a mirror WRITE request directed to the
5 mirror unified data set, the unified-sequence-number component of the supervisor local data-
storage device

determines whether a mirror WRITE request directed to the mirror unified data
set is currently pending, and

10 when a mirror WRITE request directed to the mirror unified data set is
currently pending, queues the request for a unified sequence number to a unified sequence
number request queue, and

when a mirror WRITE request directed to the mirror unified data set is not
currently pending, returns a next unified sequence number to the controller of a local data-
storage device.

15 5. The data-consistent, distributed unified data set of claim 3 wherein, when the
unified-sequence-number component of the supervisor local data-storage device receives a
request for a unified sequence number to include in a mirror WRITE request directed to the
mirror unified data set, the unified-sequence-number component of the supervisor local data-
20 storage device

determines whether a maximum number of mirror WRITE requests directed to
the mirror unified data set are currently pending, and

25 when a maximum number of mirror WRITE requests directed to the mirror
unified data set are currently pending, queues the request for a unified sequence number to a
unified sequence number request queue, and

when a maximum number of mirror WRITE requests directed to the mirror
unified data set are not currently pending, increments a count of pending mirror WRITE
requests, and returns a next unified sequence number to the controller of a local data-storage
device.

30

6. The data-consistent, distributed unified data set of claim 3

wherein the first supervisor local data-storage device is coupled to a first communications medium that links the first supervisor local data-storage device to a host computer;

5 wherein the second subordinate local data-storage device is coupled to a second communications medium that links the second subordinate local data-storage device to a host computer;

10 wherein the first supervisor local data-storage device is coupled to a third communications medium that links the first supervisor local data-storage device to a remote data-storage device to which the first supervisor local data-storage device transmits mirror WRITE requests; and

15 wherein the second subordinate local data-storage device is coupled to a fourth communications medium that links the second subordinate local data-storage device to a remote data-storage device to which the second subordinate local data-storage device transmits mirror WRITE requests.

7. The data-consistent, distributed unified data set of claim 3 wherein the first and second communications media are a single communications medium.

8. The data-consistent, distributed unified data set of claim 3 wherein the first, second, 20 third, and fourth communications media are a single communications medium.

9. The data-consistent, distributed unified data set of claim 3 wherein the first supervisor local data-storage device and the second subordinate local data-storage device are coupled to a vertical communications medium that allows the first supervisor local data- 25 storage device and the second subordinate local data-storage device to exchange unified sequence number requests and responses.

10. The data-consistent, distributed unified data set of claim 9 wherein, when a communications medium linking a particular local data-storage device to a remote data- 30 storage device fails, the particular local data-storage device transmits a mirror WRITE request to the remote data-storage device by forwarding the mirror WRITE request to the other local

data-storage device via the vertical communications medium, the other local data-storage device then forwarding the mirror WRITE request to the remote data-storage device.

11. The data-consistent, distributed unified data set of claim 3 wherein, when a remote
5 data-storage device receives a mirror WRITE request, the remote data-storage device:

extracts a unified sequence number from the mirror WRITE request;

determines whether the extracted unified sequence number is a next sequence number for the mirror unified data set;

10 when the extracted unified sequence number is a next sequence number for the mirror unified data set, carries out the WRITE request and sends the extracted unified sequence number to all other remote data-storage devices; and

when the extracted unified sequence number is not a next sequence number for the mirror unified data set, waits to execute the WRITE request until the remote data-storage device detects a preceding unified sequence number.

15

12. The data-consistent, distributed unified data set of claim 11 wherein a remote data-storage device detects a preceding unified sequence number by extracting the preceding unified sequence number from a received mirror WRITE request or by receiving the preceding unified sequence number from another remote data-storage device.

20

13. The data-consistent, distributed unified data set of claim 1 wherein, when the distributed unified data set is distributed over multiple local data storage devices, one of the local data storage devices is a supervisor data storage device and includes a unified sequence number component and the remaining local data storage devices are subordinate local data storage devices.

25
14. The data-consistent, distributed unified data set of claim 1 wherein the local data storage devices and remote data storage devices are disk arrays that include two or more disk drives.

30

15. A method for maintaining data consistency between a unified data set distributed over two or more local data storage devices, each containing a controller and each coupled to a communications medium, and a mirror unified data set stored on one or more remote data storage devices, each containing a controller and coupled to a communications medium, the
5 method comprising:

designating one of the local data storage devices as a supervisor local data storage device, and designating the remaining local data storage devices as subordinate local data storage devices;

10 executing a unified sequence number component by the controller of the supervisor local data storage device; and

prior to transmitting a mirror WRITE request directed to the mirror unified data set from a local data storage device,

15 requesting by the controller of the local data storage device a unified sequence number to associate with the mirror WRITE request from the unified sequence number component, and

including a unified sequence number received from the unified sequence number component within the mirror WRITE request.

16. The method of claim 15 further including

20 under synchronous mirroring,

when the unified-sequence-number component of the supervisor local data-storage device receives a request for a unified sequence number to include in a mirror WRITE request directed to the mirror unified data set, the unified-sequence-number component of the supervisor local data-storage device

25 determines whether a mirror WRITE request directed to the mirror unified data set is currently pending, and

when a mirror WRITE request directed to the mirror unified data set is currently pending, queues the request for a unified sequence number to a unified sequence number request queue, and

when a mirror WRITE request directed to the mirror unified data set is not currently pending, returns a next unified sequence number to the controller of a local data-storage device; and

under asynchronous mirroring,

5 when the unified-sequence-number component of the supervisor local data-storage device receives a request for a unified sequence number to include in a mirror WRITE request directed to the mirror unified data set, the unified-sequence-number component of the supervisor local data-storage device

10 determines whether a maximum number of mirror WRITE requests directed to the mirror unified data set are currently pending, and

 when a maximum number of mirror WRITE requests directed to the mirror unified data set are currently pending, queues the request for a unified sequence number to a unified sequence number request queue, and

15 when a maximum number of mirror WRITE requests directed to the mirror unified data set are not currently pending, increments a count of pending mirror WRITE requests, and returns a next unified sequence number to the controller of a local data-storage device.

17. The method of claim 15 further including

20 when receiving a mirror WRITE request by a remote data-storage device,
 extracting a unified sequence number from the mirror WRITE request;
 determining whether the extracted unified sequence number is a next sequence number for the mirror unified data set;

25 when the extracted unified sequence number is a next sequence number for the mirror unified data set, carrying out the WRITE request and sending the extracted unified sequence number to all other remote data-storage devices; and

 when the extracted unified sequence number is not a next sequence number for the mirror unified data set, waiting to execute the WRITE request until the remote data-storage device detects a preceding unified sequence number.

18. The method of claim 15 further including directly coupling the local data storage devices via one or more vertical communications media to facilitate sending of requests for unified sequence numbers and sending of unified sequence numbers.

5 19. The method of claim 15 further including, when the mirror unified data set is distributed over multiple remote data storage devices, directly coupling the remote data storage devices via one or more vertical communications media.

10 20. The method of claim 18 further including, when a communication medium between a local data storage device and a remote data storage device fails, transmitting a WRITE request from the local data storage device to the remote data storage device via a different local data storage device.